

5,745,182

1

METHOD FOR DETERMINING MOTION COMPENSATION

This is a division of application Ser. No. 07/970,046 filed Nov. 2, 1992, now U.S. Pat. No. 5,369,449.

BACKGROUND OF THE INVENTION

1. Field of the Invention

The present invention relates to a method for determining motion compensation of a moving image to be utilized in an apparatus which requires a prediction of a moving image such as an image transmission apparatus and an image apparatus.

2. Description of the Prior Art

With the progress of semiconductor technologies, methods for determining motion compensation to be utilized for a transmission of an image and a compression of an image have been widely used in many fields in recent years. Among such conventional methods for compensating for motion of a moving image, there is one method for compensating for motion of a moving image based on one piece of a reference image.

FIG. 6 is a diagram for showing the concept of the conventional method for compensating for motion of an image. Referring to FIG. 6, a moving image signal is a set of images which are sampled with an equal time interval t_0 on the time axis. For example, an NTSC signal has images sampled at every $1/60$ second for each field and a PAL signal has images sampled at every $1/50$ second for each field. When a certain object of which images are to be picked up is moving, for example, the spatial position of an object A in an M-th image is deviated from the spatial position of an object A' in an (M-1)-th image by a portion of a move of the object during a period of t_0 . Now, consider a case for predicting the M-th image from the (M-1)-th image. In order to make a determination of the M-th image with a high level of precision by compensating for motion of the object from an input image to a reference image during a time difference of t_0 , the M-th image is divided into blocks including at least one pixel, and a move of each block from the (M-1)-th image to the M-th image is detected so that a pixel value of the image at a position deviated by the portion of this move is set as a determined value. This will be explained with reference to FIG. 6. To obtain a determined value of a pixel X of the M-th image, a pixel X' at the same spatial position as the spatial position of the pixel X in the (M-1)-th image is deviated by a detected move MV of a block unit including the pixel X', so that a pixel X" is obtained. This pixel X" is then used as a determined value of the pixel X. In FIG. 6 the block is assumed to have a size of 3×3 .

When a signal is an interlace signal, there are many alternative cases considered for predicting compensation for motion of an image. For example, either a frame or a field is used for the image, and a frame is used for a reference image and a field is used for an input image, etc. The basic principle is as explained with reference to FIG. 6 above. As one of the examples of the above method for predicting motion compensation, there is Recommendation 723, "Transmission of component-coded digital television signals for contribution-quality at the third hierarchical level of CCITT Recommendation G.702" which was standardized by the CCITT (Commission Mixte CCIR/CCITT pour les Transmissions Televisuelles et Sonores 3). In this recommendation, a determination of motion compensation between frames and a determination of motion compensa-

2

tion between fields are suitably changed over between the two cases. As described above, according to the conventional method for determining motion compensation of an image, a determination is made by compensating for motion of the image based on detected motion of the image. Therefore, the conventional predicting method can predict motion compensation with a high level of precision even if an image is a moving image including movement.

The above-described conventional method for determining motion compensation, however, has problems that it is not possible to accurately determine motion compensation and that, even if it is possible to correctly determination of motion compensation, the image density of an image to be referred to becomes the image density of a reference image, which makes it impossible to make prediction at a higher level of precision.

For example, in the case of determining motion compensation by using an interlace signal as a frame and generating a block from this frame, frames are combined together to compensate motion of an image by disregarding a difference in sampling positions, due to a time difference, between two fields within a frame. Accordingly, when correct sampling positions of the fields are considered, there is such a case that motion compensated in the first field and motion compensated in the second field do not coincide with each other. An example of this case is shown in FIGS. 7A to 7C. Referring to FIGS. 7A to 7C, an input signal is an interlace signal (FIG. 7A). Interlace signals are combined together in a frame to determine motion compensation. When a vertical component of a motion detected now is 1, the first field of the M-th frame is predicted from the second field of the (M-1)-th frame and the second field of the M-th frame is predicted from the first field of the (M-1)-th frame, as shown in FIG. 7B. Moves in the correct field positions is shown in FIG. 7C. As is clear from FIG. 7C, the motion for effecting compensation in the first field of the M-th frame do not coincide with the moves for effecting compensation in the second field of the M-th frame. As explained above, when motion compensation of an image is made by handling an interlace image as a frame, the motion for effecting compensation are different between the first field and the second field. In a vector in which this phenomenon occurs, there is a problem that the precision of the level of prediction is deteriorated.

Next, consider a case of determining motion compensation of an image as an image of a correct position without disregarding a time difference of sampling between images as described above. As examples of this case, there is a case where motion compensation is determined for an interlace signal by generating a block from a field, and a case where motion compensation is determined for a noninterlace signal. In the above cases, motion compensation is predicted by using an image at a position of a correct time. Therefore, there arises no such problem which occurs in the case of determined motion compensation by generating a block from a frame of the interlace signal as described above. However, in this case, motion compensation is determined from one piece of reference image and the pixel density of an image to be referred to becomes the pixel density of the reference image, so that there is a limit to carrying out a determination of motion compensation at a higher level of precision. FIG. 8 shows a case of determined move compensation by generating a block from a field for an input of an interlace signal. In this case, determination of motion compensation is carried out by using a field image as a reference image. Therefore, when a motion vector is 0 there is no sampling point at a position necessary for making a determination on the reference image and, accordingly, a

5,745,182

3

pixel value, or a determined value, must be calculated by interpolation within the field, as shown in FIG. 8, for example. As compared with the case for compensating motion by generating a block based on a pixel value within a frame, the case for compensating motion based on the field has a pixel density in a vertical direction which is half of the pixel density in the case of compensating a move based on a frame. Thus, there is a limit to carrying out a determination of motion compensation at a high level of precision when motion compensation is carried out based on a field. This problem also arises when motion compensation is carried out by using a non-interlace signal as an input. In both cases, the pixel density of the image to be referred to becomes the pixel density of the reference image, and there is a limit to carrying out a determination of motion compensation at a higher level of precision.

SUMMARY OF THE INVENTION

With a view to eliminating the above-described problems of the prior-art technique, it is an object of the present invention to provide a method for determining motion compensation with a very high level of precision by utilizing a plurality of pieces of reference images.

In order to achieve the above-described object of the present invention, the method of the present invention determines motion compensation of an input image based on a motion vector of a reference image from an original position of the reference image to a position of the reference image sampled at a first set time, and the method includes calculating a motion vector of an input image by calculating a motion at a second set time of a block unit which is a part of the input image and also consists of a plurality of pixels, and calculating a motion vector of the reference image by calculating a move at the first set time of a block unit which is a part of the reference image and also consists of a plurality of pixels, to thereby calculate motion compensation of the input image at a desired set time both from the motion vector of the input image and from the motion vector of the reference image.

Also, the method of the present invention determines motion compensation of a plurality of pieces of input images based on a motion vector of a reference image from an original position of the reference image to a position of the reference image sampled at a first set time, and the method includes calculating motion vectors of input images by calculating motion at a second set time of block units, each block forming a part of each input image and also consisting of a plurality of pixels, and a unit for calculating a motion vector of the reference image by calculating a motion at the first set time of a block unit which is a part of the reference image and also consists of a plurality of pixels, regarding these motion vectors of the input images to be the same, to thereby calculate motion compensation of the input images at a desired set time both from the motion vectors of the input images and the motion vector of the reference image.

Therefore, according to the present invention, a time position of a reference image is compensated by using a certain motion vector depending on the need so that a plurality of pieces of reference images sampled at different times according to the motion of a block unit, including at least one pixel, which is detected at a certain time interval, become images of the input image at the above time intervals from the position of the input image. Thus, it is possible to obtain a plurality of pieces of images at positions of the above-described time intervals from the position of the input image. By combining these images together, it is possible to

4

obtain a reference image of high pixel density. Based on this reference image of high pixel density, a pixel value at a position compensated by the detected motion portion is calculated and this is used as a determined value. Accordingly, it is possible to determine motion compensation at a very high level of precision.

Further, according to the present invention, a vector for carrying out compensation of a time position of the reference image can be calculated from motion of the image detected at a certain time interval, so that it is not necessary to detect again the motion vector for correcting the time, and motion compensation at a high level of precision can be ensured.

Further, by using an interlace signal as an input signal and using two fields in a certain frame for a reference image, it becomes possible to suitably apply the above method for determining motion compensation to a frame image, thus ensuring a determination at a high precision level of motion compensation based on a frame.

Further, since the same value is used for a block of each input image, of which whole or part of spatial position of each block is superposed, among blocks of a plurality of pieces of input images, as motion detected at a certain time interval in a block unit including at least one pixel, it is not necessary to carry out motion detection a plurality of times for many blocks of the plurality of input images, thus ensuring a determination of motion compensation at a high level of precision.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 is a diagram for explaining the first embodiment of the present invention;

FIGS. 2A and 2B are diagrams for explaining the second embodiment of the present invention;

FIG. 3 is a diagram for explaining a block for determining motion compensation using a frame as a base;

FIGS. 4A and 4B are diagrams for explaining the third embodiment of the present invention;

FIG. 5 is a diagram for explaining the fourth embodiment of the present invention;

FIG. 6 is a conceptional diagram for showing the conventional method for determining motion compensation;

FIG. 7A to FIG. 7C are diagrams for explaining problems of the conventional method for determining motion compensation between frames; and

FIG. 8 is a diagram for explaining the conventional method for determining motion compensation between fields.

DESCRIPTION OF THE PREFERRED EMBODIMENTS

FIG. 1 is a diagram for explaining a first embodiment of the present invention. FIG. 1 assumes the determination of motion compensation based on an interlace signal as an input signal so that a block is generated from an image within a field and a field image is used as a base. In this case, the input image is in the M-th field and a reference image is in both the (M-1)-th field and the (M-2)-th field. Now, assume that a motion vector (MV) for predicting move compensation of a certain block is to be detected in a two-field instance, that is, between the M-th field and the (M-2)-th field. To simplify the explanation, of the detected moves, only motion in the vertical direction will be considered, and a pixel value is expressed as $a(x, y)$. In this

5,745,182

5

case, x represents a field number and y represents a line number. Line numbers are placed in the order of 1, 2, . . . , starting from the bottom in a line interval of the frame. A position of each pixel in a vertical direction is expressed always in the unit of a frame line.

Now, consider a case for obtaining a determined value of $(M, 1)$. When a vertical component of a detected MV is 1, a determined value of motion compensation of a $(M, 1)$ becomes a pixel value at a position of $(M-2, 2)$. Next, a time position is corrected so that the position of the reference image in the $(M-1)$ -th field becomes an image at the position of the $(M-2)$ -th field. A vector for correcting this time position is set as MV adj (adjustment vector). In calculating the MV adj from the MV, the following relationship can be easily drawn when it is assumed, for example, that motion from the $(M-2)$ -th field to the M -th field is constant:

$$MV \text{ adj} = -MV/2$$

Therefore, when the vertical component of the MV is 1, the vertical component of the MV adj becomes -0.5 . When this is applied to the positional correction for correcting the positions of the $(M-1)$ -th field to the positions of the $(M-2)$ -th field, a $(M-1, 2)$ in the $(M-1)$ -th field is corrected to the position of $(M-2, 2.5)$ in the $(M-2)$ -th field, as shown in FIG. 1. After the above operation, the position of $(M-2, 2)$ which is the determined value of motion compensation of the $(M, 1)$ is obtained from pixel values of the $(M-2)$ -th field and the $(M-1)$ -th field which is the result of a time positional correction. In this case, when a necessary pixel value is to be obtained by a weighted average, which is inversely proportional to a distance, from near-by pixel values, for example, the pixel value at the position of the $(M-2, 2)$, that is the determined value of motion compensation of the $(M, 1)$, is obtained by the following expression:

$$a(M-2, 1) \cdot 3 + 2 \cdot a(M-1, 2) \cdot 3$$

Although the above explains a determination of motion compensation taking account of only vertical components, a similar operation is also applied to the case for determining motion compensation for an image having both vertical and horizontal components.

As described above, according to the first embodiment of the present invention, a time position of a reference image is corrected by using a motion vector as required so that a plurality of pieces of reference images sampled at different times become an image at a time which is positioned within a certain time interval from the position of the input image, according to detected motion at the above time interval of a block unit including at least one pixel. Therefore, it is possible to obtain a plurality of pieces of images at positions which are separated from the input image by the above time intervals. A reference image with high pixel density is obtained by combining these pieces of images, and a pixel value at a position which compensates the detected motion is calculated by using the reference image of high pixel density, and the calculated pixel value is used as a determined value. As a result, there is an effect that it is possible to carry out a determination of motion compensation of an image at a very high precision level.

The assumptions used in the above description of the first embodiment are flexible and the following alternative assumptions can also be accepted. Instead of assuming two images for a reference image, any plurality of number of

6

pieces of images can also be used. Instead of assuming the positions of the plurality of reference images as the previous image and its previous image, these positions can also be set flexibly. Instead of assuming constant motion of an image for calculating the MV adj from the MV, any assumption of motion can also be used according to a certain rule for calculating the MV adj from the MV. Instead of using a weighted average which is inverse proportional to a distance from near-by pixel values to obtain a pixel value at a necessary position, a coefficient of a low-pass filter, for example, can also be used to obtain the pixel value at a necessary position. Instead of using interpolation, for obtaining a necessary pixel value, extrapolation, for example, can also be used to obtain the necessary pixel value.

Further, although the vector M adj for correcting positions of a plurality of pieces of reference images is calculated from a motion vector MV which is detected in a block unit in the above-described first embodiment of the present invention, the MV can also be detected independently between the $(M-1)$ -th field and the $(M-2)$ -th field. In this case, by an independent detection of the MV, the time correction can be effected based on a more accurate motion of the image. Thus, this method has an effect that a determination of motion compensation of an image at a higher level of precision can be ensured.

Further, although an interlace signal is used as an input signal and a field image is assumed as each image in the above-described first embodiment of the present invention, a non-interlace image can also be used, with an effect that a determination of motion compensation of an image can be ensured at a higher level of precision, for the same reason as explained above.

Next, a second embodiment of the present invention for determining motion compensation of an image by using an interlace signal as an input signal and using a frame as an input, will be explained. FIGS. 2A and 2B are diagrams for explaining the second embodiment of the present inventions. Referring to FIGS. 2A and 2B, a reference image is in two fields of the previous frame, that is, the $(M-1)$ -th field and the $(M-2)$ -th field, and an input image is in two fields of the current frame, that is, the M -th field and the $(M+1)$ -th field. In FIGS. 2A and 2B, it is assumed that a motion vector for determining motion compensation of a certain block is obtained within an input image and a reference frame and between fields of the same phase as that of the M -th field, for each two fields of the current frame. The motion vector for predicting move compensation of the M -th field is expressed as $MV(M)$ and the motion vector for the $(M+1)$ -th field is expressed as $MV(M+1)$. A pixel value of each pixel position is expressed in the same manner as that of the first embodiment. To simplify the explanation of FIGS. 2A and 2B, of the detected motion, only the move in the vertical direction will be considered.

In FIGS. 2A and 2B, a determination of a pixel within the M -th field is carried out by using images in the $(M-1)$ -th field and the $(M-2)$ -th field, in exactly the same operation as that of the first embodiment of the present invention. For example, when the vertical component of the $MV(M)$ is 1.5, the determined value of $a(M, 1)$ becomes the pixel value at the position of $(M-2, 2.5)$ according to the operation which is the same as that of the first embodiment, and this value is obtained by the following expression:

$$a(M-2, 1) \cdot 7 + 6 \cdot a(M-1, 2) \cdot 7$$

Similarly, in the same manner as that of the determination of a pixel within the M -th field, a pixel within the $(M+1)$ -th

5,745,182

7

field is determined from the two fields of the reference frame, that is, the (M-1)-th and (M-2)-th fields. The method of determination in this case is the same as the method for determining a pixel within the M-th field, except that the (M-2)-th field needs to be corrected to the position of the (M-1)-th field. A vector for correcting this time position is set as MV adj (M+1). In calculating the MV adj (M+1) from the MV(M+1), when the motion from the (M-2)-th field to the (M+1)-th field, for example, are constant, the following relationship can be obtained easily:

$$MV \text{ adj } (M+1) = MV(M+1)/2$$

Therefore when the vertical component of the MV(M+1) is 1, the vertical component of the MV adj (M+1) becomes 0.5. As shown in FIGS. 2A and 2B, when the (M-2)-th field is positionally corrected to the position of the (M-1)-th field, the a(M-2, 3) is positionally corrected to the position of the (M-1, 2.5). After the above operation, the position of the (M-1, 3) which is a determined value of the a(M+1, 2) is obtained from the pixel value of the (M-1)-th field and the pixel value of the (M-2)-th field of which time position has been corrected. When a necessary pixel value is to be obtained by a weighted average, inversely proportional to a distance, from near-by pixel values, for example, a determined value at the position of the (M-1, 3), that is, a determined value of motion compensation of the a(M+1, 2), is obtained by the following expression:

$$a(M-1, 4)/3 + 2 * a(M-2, 3)/3$$

Although the above explains the determination of motion compensation of an image for only the vertical component, a similar operation is also applied to the case for determining motion compensation of an image having both vertical and horizontal components.

As described above, according to the second embodiment of the present invention, the above-described determination of motion compensation can be applied to a frame image, by using an interlace signal as an input signal and by setting a reference image in two fields of a certain frame. As a result, there is an effect that it is possible to determine motion compensation of an image at a high level of precision by using a frame as a base.

The above-described second embodiment also has a flexibility in the assumptions used, in the same manner as that of the first embodiment. For example, the number of reference frames, the positions of the reference frames, the assumptions for obtaining the MV adj (M) or MV adj (M+1) from either the MV(M) or the MV(M+1), the calculation method for obtaining a pixel value at a necessary position, and either interpolation or extrapolation, can also be selected freely. Further, although it is assumed in the present embodiment that a move vector for determining motion compensation can be obtained within the input image and reference frame and between fields of the same phase as that of the input image, it is also possible to obtain the motion vector between fields of opposite phases, in the same manner of operation, with the similar effect. Further, when a position correction vector is obtained independent of a detection motion vector, in the same manner as that of the first embodiment, there is an effect that it is possible to determine motion compensation at a higher level of precision.

Next, as a third embodiment of the present invention, another method for determining motion compensation for an interlace input signal based on a frame unit will be

8

explained. FIG. 3 and FIGS. 4A and 4B are diagrams for explaining the third embodiment of the present invention. Referring to FIG. 3, a reference image is in the (N-1)-th frame, that is, the (M-2) and (M-1) fields, and an input image is in the N-th frame, that is, the M-th and (M+1)-th fields. Now assume that a block for carrying out motion compensation is being generated from the frame. Assume that a motion vector MV is to be obtained in the block unit generated from the pixels of the N-th frame, from the (N-1)-th frame. The status of the block in this case is shown in FIG. 3. In terms of the method for determining motion compensation based on a field, the following method can be considered. The reference image is in the two fields of the (N-1)-th frame and the input image is in the two fields of the N-th frame. The detecting interval of the MV is the two-field interval. However, the pixels included in the above block have the same motion vector MV for both the pixels in the M-th field and the pixels in the (M+1)-th field.

In other words, in the case of the third embodiment, motion vectors to be used for pixels within a block generated from the above frame take the same value regardless of whether the pixels belong to the M-th field or the (M+1)-th field. The other operations become the same as those of the second embodiment. FIGS. 4A and 4B show the case that the vertical component of the MV is 1. Although the above explains the case of determining motion compensation of an image for only the vertical component, a similar operation is also carried out for the case of determining motion vector of an image having both vertical and horizontal components.

As described above, according to the third embodiment of the present invention, the same motion vector is used for pixels in the two input fields positioned within a predetermined spacial area such as a block generated by the frame. Accordingly, it is not necessary to detect motion vectors for each field according to this method, which also has an effect that it is possible to determine motion compensation at a high level of precision.

The above-described third embodiment also has a flexibility in the assumptions used, similar to the case of the second embodiment. For example, the number of reference frames, the positions of the reference frames, the assumption for obtaining the MV adj from the MV, the calculation method for obtaining a pixel value at a necessary position, and whether interpolation or extrapolation is to be used, can all be selected freely. Although description has been made of the case for determining motion compensation based on a frame as a unit in the present embodiment, it is needless to mention that there is no change in the effect of determination if the determination is carried out based on a field as shown in the first embodiment or if the determination is carried out based on a noninterlace image. Further, when a block to be used for having the same value of the motion vector is selected from among blocks of a plurality of pieces of input images in such a way that the block selected is a block of each input image of which part or whole of the spatial position superposes with those of the other blocks selected, there is no change in the effect of prediction. Further, similar to the case of the second embodiment, by obtaining a position correction vector independent of a detection motion vector, there is an effect that it is possible to determine motion compensation at a higher level of precision.

FIG. 5 is a diagram for explaining a fourth embodiment of the present invention. The fourth embodiment takes the same assumptions as those of the first embodiment, and an interlace signal is used as an input signal, an input image is in the M-th field and a reference image is in both the (M-1)-th field and the (M-2)-th field. Assume in FIG. 5 that